# Basis set convergence of the coupled-cluster correction, δMP2CCSD(T): Best practices for benchmarking non-covalent interactions and the attendant revision of the S22, NBC10, HBC6, and HSG databases

Michael S. Marshall, Lori A. Burns, and C. David Sherrill

Citation: J. Chem. Phys. **135**, 194102 (2011); doi: 10.1063/1.3659142
View online: http://dx.doi.org/10.1063/1.3659142
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v135/i19
Published by the American Institute of Physics.

# Basis set convergence of the coupled-cluster correction, $\delta_{MP2}^{CCSD(T)}$: Best practices for benchmarking non-covalent interactions and the attendant revision of the S22, NBC10, HBC6, and HSG databases

Michael S. Marshall, Lori A. Burns, and C. David Sherrill[a)]
*Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA*

(Received 19 August 2011; accepted 19 October 2011; published online 15 November 2011)

In benchmark-quality studies of non-covalent interactions, it is common to estimate interaction energies at the complete basis set (CBS) coupled-cluster through perturbative triples [CCSD(T)] level of theory by adding to CBS second-order perturbation theory (MP2) a "coupled-cluster correction," $\delta_{MP2}^{CCSD(T)}$, evaluated in a modest basis set. This work illustrates that commonly used basis sets such as 6-31G*(0.25) can yield large, even wrongly signed, errors for $\delta_{MP2}^{CCSD(T)}$ that vary significantly by binding motif. Double-$\zeta$ basis sets show more reliable results when used with explicitly correlated methods to form a $\delta_{MP2-F12}^{CCSD(T^*)-F12}$ correction, yielding a mean absolute deviation of 0.11 kcal mol$^{-1}$ for the S22 test set. Examining the coupled-cluster correction for basis sets up to sextuple-$\zeta$ in quality reveals that $\delta_{MP2}^{CCSD(T)}$ converges monotonically only beyond a turning point at triple-$\zeta$ or quadruple-$\zeta$ quality. In consequence, CBS extrapolation of $\delta_{MP2}^{CCSD(T)}$ corrections before the turning point, generally CBS (aug-cc-pVDZ,aug-cc-pVTZ), are found to be unreliable and often inferior to aug-cc-pVTZ alone, especially for hydrogen-bonding systems. Using the findings of this paper, we revise some recent benchmarks for non-covalent interactions, namely the S22, NBC10, HBC6, and HSG test sets. The maximum differences in the revised benchmarks are 0.080, 0.060, 0.257, and 0.102 kcal mol$^{-1}$, respectively. © *2011 American Institute of Physics*. [doi:10.1063/1.3659142]

## I. INTRODUCTION

High-accuracy benchmark data sets have become a cornerstone for testing new theories, basis sets, and approximations in computational chemistry. As the field becomes more reliant on these high-quality test sets, a better understanding of their underlying errors is required. One common approach for obtaining benchmark quality reference data is through focal-point analysis.[1,2] In the context of non-covalent interactions, focal-point analysis is often used to estimate coupled-cluster theory through perturbative triple excitations in a large basis set [est. CCSD(T)/large]:

$$E_{CCSD(T)}^{large} \approx E_{MP2}^{large} + \delta_{MP2}^{CCSD(T)}, \quad (1)$$

$$\delta_{MP2}^{CCSD(T)} = E_{CCSD(T)}^{small} - E_{MP2}^{small}. \quad (2)$$

The subscripts and superscripts on energies refer to the method and basis set, respectively. The coupled-cluster correction $\delta_{MP2}^{CCSD(T)}$ is also referred to as ΔCCSD(T) in the literature. The "large" basis set is often a complete basis set (CBS) extrapolation, whereas the "small" basis set used for $\delta_{MP2}^{CCSD(T)}$ is usually a single basis set, though it may also be a basis set extrapolation.[3] As has been noted,[4–8] this focal-point approach works well because even though the convergence rates of CCSD(T) and MP2 correlation energies are slow with respect to basis set size, the rate of convergence of the difference [CCSD(T)–MP2] is much faster.[9] Thus, a much smaller basis set may be used for the $\delta_{MP2}^{CCSD(T)}$ correction than for the underlying SCF and MP2 computations. This approach has been applied to obtain several non-covalent interaction benchmark energies.[4,10–12] One of the more common small basis sets used is 6-31G*(0.25), particularly for benchmarking non-covalent interactions in bio-molecules. This modified Pople basis set is formed by replacing the usual exponent for the *d* polarization functions with a more diffuse exponent ($\alpha_d = 0.25$) better able to describe non-bonding interactions.[13–16] While this approach has been used widely, there now exist several papers showing that $\delta_{MP2}^{CCSD(T)}$ corrections with modest basis sets (e.g., 6-31G*, cc-pVDZ, and aug-cc-pVDZ) often lead to inaccurate interaction energies. One such paper by Boese *et al.*[17] reports interaction energy errors of 10% for neutral hydrogen-bonded complexes using 6-31G*(0.25) and claims that using such small basis sets "does more harm than good" (because too small a basis set can yield the wrong sign for the $\delta_{MP2}^{CCSD(T)}$ correction). A large error is also seen in the work of Min *et al.* on benzene · $Na^+$ (Ref. 18). They report a $\delta_{MP2}^{CCSD(T)}$ correction increasing in magnitude from −0.13 kcal mol$^{-1}$ in the aug-cc-pVDZ basis to −1.16 kcal mol$^{-1}$ in the aug-cc-pVTZ basis. Pitoňák *et al.*[8] recently studied the effect of small basis sets on stacked adenine · thymine. They find that small Pople basis sets have errors less than 10%–20% in the $\delta_{MP2}^{CCSD(T)}$ correction. While this nucleobase test system is interesting, it is not representative of difficult cases in which the $\delta_{MP2}^{CCSD(T)}$ correction is a significant portion of the binding energy. In stacked adenine · thymine, $\delta_{MP2}^{CCSD(T)}$ is 3.18 kcal mol$^{-1}$ compared to a total interaction energy of −11.66 kcal mol$^{-1}$;

a)Electronic mail: sherrill@gatech.edu.

 

whereas, the more difficult cases are systems like the parallel-displaced benzene dimer where the $\delta_{MP2}^{CCSD(T)}$ correction is 1.96 kcal mol$^{-1}$ compared to a total interaction energy of only −2.67 kcal mol$^{-1}$ (Ref. 3). When the $\delta_{MP2}^{CCSD(T)}$ correction is large compared to the interaction energy, a significant error in this quantity can result in a large relative error in the total binding of the complex. This is yet another reason why the benzene dimer has been examined in such detail.[3,4,10,19–23] In particular, a key study by Janowski and Pulay[21] demonstrated that even the reasonably good aug-cc-pVDZ basis set is not quite sufficient to converge $\delta_{MP2}^{CCSD(T)}$ within 0.1 kcal mol$^{-1}$ of its true value for the benzene dimer. Given that some approximate methods for non-covalent interactions are now achieving mean absolute deviations (MADs) of only few tenths of one kcal mol$^{-1}$ (Refs. 24 and 25), it is important to begin considering how the remaining errors in the benchmark interaction energies can be reduced to only a few hundredths of one kcal mol$^{-1}$. In many cases, the $\delta_{MP2}^{CCSD(T)}$ term appears to be the largest remaining source of error, and hence it is the focus of the present study. In particular, we examine the basis set convergence of this correction for several small van der Waals dimers, and from this work we present revised benchmark interaction energies for several databases of non-covalent interactions.

While estimates of CCSD(T)/CBS interaction energies using a relatively small basis set to evaluate $\delta_{MP2}^{CCSD(T)}$ can be quite accurate in many cases, the quality of the procedure is not necessarily consistent across binding motifs for non-covalent interactions. This study examines the error incurred by using double-$\zeta$ basis set $\delta_{MP2}^{CCSD(T)}$ corrections for complexes that are hydrogen-bonded, dispersion-bound, or of mixed character. Particularly, we focus on the S22 benchmark set,[11] which features diverse types of non-bonded interactions over a wide range of system sizes, from water dimer (six atoms) to adenine · thymine complexes (30 atoms).

This work also investigates the use of explicitly correlated wavefunctions to obtain better $\delta_{MP2}^{CCSD(T)}$ corrections at reduced computational cost. Analogous to Eqs. (1) and (2),

$$E_{CCSD(T^*)-F12}^{large} \approx E_{MP2-F12}^{large} + \delta_{MP2-F12}^{CCSD(T^*)-F12}, \quad (3)$$

$$\delta_{MP2-F12}^{CCSD(T^*)-F12} = E_{CCSD(T^*)-F12}^{small} - E_{MP2-F12}^{small}. \quad (4)$$

Explicitly correlated wavefunctions have been shown to produce accurate energies using relatively small basis sets.[26–31] This work evaluates whether this same approach can more quickly converge the coupled-cluster correction.

## II. THEORETICAL APPROACH

### A. Notation

The Dunning basis sets aug-cc-pVNZ (N = D,T,Q,5,6) are herein referred to as aNZ. The heavy-aug-cc-pVNZ (N = D,T,Q,5,6), which is aug-cc-pVNZ on the non-hydrogen atoms and cc-pVNZ on the hydrogens, is herein referred to as haNZ. It is also worth noting here that none of the interaction energies presented in this work are at fully optimized minima, but instead at fixed, near-equilibrium geometries defined by the corresponding original papers. All interaction energies are counterpoise corrected for basis set superposition error (BSSE) with the scheme outlined by Boys and Bernardi.[32]

### B. Convergence of complete basis set extrapolated $\delta_{MP2}^{CCSD(T)}$ corrections

A better understanding is required of the convergence trends of $\delta_{MP2}^{CCSD(T)}$ corrections with respect to basis size for both individual basis sets (e.g., aNZ, N = D,T,Q,5,6) and 2-point Helgaker extrapolations[33] (e.g., CBS(aNZ,aMZ), NM = DT, TQ, Q5, 56). To accomplish this, $\delta_{MP2}^{CCSD(T)}$ quantities are computed for each of the aforementioned basis sets for several of the smallest members of the S22 test set.

### C. Revision of benchmark databases

In accordance with the conclusions of this study (discussed below), the reference interaction energies for the S22,[11] NBC10,[3,34,35] HBC6,[36] and HSG[12] databases have been revised, with geometries remaining unchanged. Benchmark values for the latter two have been computed as a sum of the HF/aQZ energy, the two-point (aTZ and aQZ) Helgaker CBS extrapolation[33] of the MP2 correlation energy, and the $\delta_{MP2}^{CCSD(T)}$ correction with the aTZ basis set (haTZ for HSG). Similarly, for the NBC10 test set, previous benchmark energies employing (h)aDZ/(h)aTZ extrapolations of $\delta_{MP2}^{CCSD(T)}$ have been updated to use simple (h)aTZ values of $\delta_{MP2}^{CCSD(T)}$. The databases at this revision level will be denoted as NBC10A, HBC6A, and HSG-A.

For the S22 test set, revisions of the original[11] benchmark energies have already been published. A recent paper by Takatani *et al.*[37] contributed high-quality estimates of the $\delta_{MP2}^{CCSD(T)}$ correction to define the S22A binding energies. The $\delta_{MP2}^{CCSD(T)}$ corrections were extrapolated to the CBS limit with an aDZ,aTZ 2-point Helgaker extrapolation[33] of the MP2 and CCSD(T) correlation energies. Podeszwa *et al.* independently revised binding energies for the S22 dimers using larger basis MP2/CBS energies and single-basis $\delta_{MP2}^{CCSD(T)}$ corrections with mid-bond functions.[38] These two studies agree within 0.044 kcal mol$^{-1}$ averaged across the entire set, or within 0.029 kcal mol$^{-1}$ if the adenine · thymine complexes (numbered 7 and 15) are dropped. Recently, our group has analyzed the differences in these two benchmarks and concluded that aDZ,aTZ CBS extrapolated $\delta_{MP2}^{CCSD(T)}$ corrections typically slightly overestimate $\delta_{MP2}^{CCSD(T)}$ corrections for hydrogen-bonded complexes. A detailed study of this is presented in Sec. III B. Additionally, we have performed new CCSD(T) computations using larger basis sets when feasible. The best quality results from among the literature values and new computations have been judiciously selected to form the S22B benchmark set. The S22 set conveniently partitions complexes by binding type into hydrogen bonding, dispersion-dominated, and mixed influence categories. This grouping allows one to see if particular methods struggle for certain kinds of non-covalent interactions. Interaction energy decompositions via DFT-SAPT by Grafová *et al.*[24] and SAPT2+(3)/aTZ results by Hohenstein and Sherrill[39] have shown that the original, intuitive assignments of the S22 complexes to

binding-type subgroups were not wholly consistent. In this work, we use the grouping suggested by SAPT2+(3) data which moves stacked adenine · thymine (15) and uracil (13) to the mixed influence subset and T-shaped benzene dimer (20) to the dispersion-dominated subset. We note that the "hydrogen bonding" group might more precisely be designated as "electrostatically dominated," as the electrostatic character is what we confirmed by SAPT analysis. Some members of this group, such as $NH_3$ dimer, may not necessarily fit the latest International Union of Pure and Applied Chemistry definition of hydrogen bonding.[40,41]

All subsequent discussions employ the revised S22B, NBC10A, HBC6A, and HSG-A interaction energies as benchmarks.

### D. Small basis set $\delta_{MP2}^{CCSD(T)}$ corrections

For each of the complexes in the S22 test set, we report $\delta_{MP2}^{CCSD(T)}$ corrections (using Eq. (2)) for the following basis small sets: 6-31G*, 6-31G*(0.25), 6-31G**(0.25,0.15), cc-pVDZ, and aug-cc-pVDZ (aDZ). The number in parentheses indicates the non-standard exponent on the polarization functions. These modified Pople basis sets were chosen because they are commonly used in the literature for computing $\delta_{MP2}^{CCSD(T)}$.[42–49] Results are compared to the best currently available values of $\delta_{MP2}^{CCSD(T)}$ (several of which are revised in this work). Mean absolute deviation as well as mean absolute percent deviation (MAPD, with respect to the total interaction energy) are reported for each basis set. Considering the minimal role of core correlation towards overall interaction energies,[17] the frozen-core approximation was employed for all computations. All total energies were converged to $10^{-9}$ hartree.

### E. Explicitly correlated $\delta_{MP2-F12}^{CCSD(T^*)-F12}$ corrections

Using Eqs. (3) and (4), we examine the performance of explicitly correlated F12 methods for computing $\delta_{MP2}^{CCSD(T)}$ corrections. Some technical aspects of CCSD(T)-F12 require elaboration, the first of which is how to handle the perturbative triples correction. This work follows the approach of Marchetti *et al.*[50] whereby the triples correction is scaled by the ratio of MP2 correlation energy and MP2-F12 correlation energy:

$$E_{(T^*)}^{corr} = E_{(T)}^{corr} \frac{E_{MP2-F12}}{E_{MP2}}. \quad (5)$$

This procedure leads to a better triples correction, but if it is done independently for the dimer and each monomer, size-consistency is lost, as pointed out by Marchetti *et al.*[50] To retain size-consistency while computing an interaction energy, the scaling factor must be kept consistent for each complex. Any of the three scaling factors (dimer, monomer A, monomer B) could be chosen, though the dimer is the most common choice. Methods that employ a single scaling factor for each of the three computations are herein referred to as CCSD(T**)-F12. CCSD(T*)-F12 (with one asterisk) here designates independently scaled triples corrections. The second issue that must be considered is the choice of F12 ansatz. For CCSD-F12, we present both F12a and F12b (Refs. 30 and 31). For MP2-F12, we present only MP2-F12/3C(FIX) as the MP2 reference.[29]

## III. RESULTS AND DISCUSSION

### A. Complete basis set extrapolated $\delta_{MP2}^{CCSD(T)}$ corrections

First, it is worthwhile to establish the best possible benchmark values for $\delta_{MP2}^{CCSD(T)}$. In previous work, we presented a basis-set-consistent revision of interaction energies for the S22 test set, which we designated S22A.[37] Based on limited comparisons where extrapolated CCSD(T)/CBS(aTZ,aQZ) interaction energies were then available, it appeared that using CBS(aDZ,aTZ) $\delta_{MP2}^{CCSD(T)}$ corrections was very similar to and in some cases slightly preferable to using aTZ $\delta_{MP2}^{CCSD(T)}$ corrections. However, in the present work, we find that this previous picture is somewhat misleading because—surprisingly—the CCSD(T)/CBS(aTZ,aQZ) interaction energies are themselves not always fully converged.

Figure 1 illustrates the slow, non-monotonic convergence characteristic of the $\delta_{MP2}^{CCSD(T)}$ term, which often results in significant errors for estimates involving double-$\zeta$ values (i.e., aDZ and CBS(aDZ,aTZ)). The $\delta_{MP2}^{CCSD(T)}$ correction for the water dimer in Figure 1(a) grows with increasing basis set until aQZ, but decreases with the a5Z and a6Z basis sets. To achieve a nearly-converged CBS $\delta_{MP2}^{CCSD(T)}$ correction, one would have to acquire a CBS(aQZ,a5Z) extrapolated estimate. Because of the scaling of CCSD(T), employing basis sets of this size is infeasible for any but the smallest complexes. From Figure 1(a), we also note how CBS(aDZ,aTZ) significantly overestimates the best available [a6Z or CBS(a5Z,a6Z)] estimates (by ~0.04 kcal mol$^{-1}$ out of a ~0.03 kcal mol$^{-1}$ correction). This is due not only to the poor quality of the aDZ basis set, but also to the fact that $\delta_{MP2}^{CCSD(T)}$ terms do not converge monotonically with increasing basis set size.

We apply this same analysis to the double hydrogen-bonded system formic acid dimer, shown in Figure 1(b). Again, we note that the $\delta_{MP2}^{CCSD(T)}$ correction does not converge monotonically and this causes the CBS(aDZ,aTZ) estimate to overshoot by −0.07 kcal mol$^{-1}$ compared to CBS(aQZ,a5Z). Figure 1(c) demonstrates the same trends for formamide dimer. In this case, the aDZ basis provides the correct sign for $\delta_{MP2}^{CCSD(T)}$, but it is only −0.07 kcal mol$^{-1}$ compared to a best estimate of −0.27 kcal mol$^{-1}$. The CBS(aDZ,aTZ) extrapolation overshoots the best estimate by 0.07 kcal mol$^{-1}$.

While these errors are not very large, they are undesirable and possible to avoid in general without costlier computations. Based on the water dimer, formic acid dimer, and formamide dimer test cases, it appears that the aTZ basis set provides a nice Pauling point for the $\delta_{MP2}^{CCSD(T)}$ correction for hydrogen-bonded systems. For two out of three cases, the aDZ basis provided the wrong sign for $\delta_{MP2}^{CCSD(T)}$, and in the other case, it achieved only one-fourth of the true value. Hence, for small hydrogen-bonded systems, we urge caution in using double-$\zeta$ basis sets for $\delta_{MP2}^{CCSD(T)}$ corrections for benchmark quality work. For somewhat larger complexes, contributions from dispersion forces will grow, and

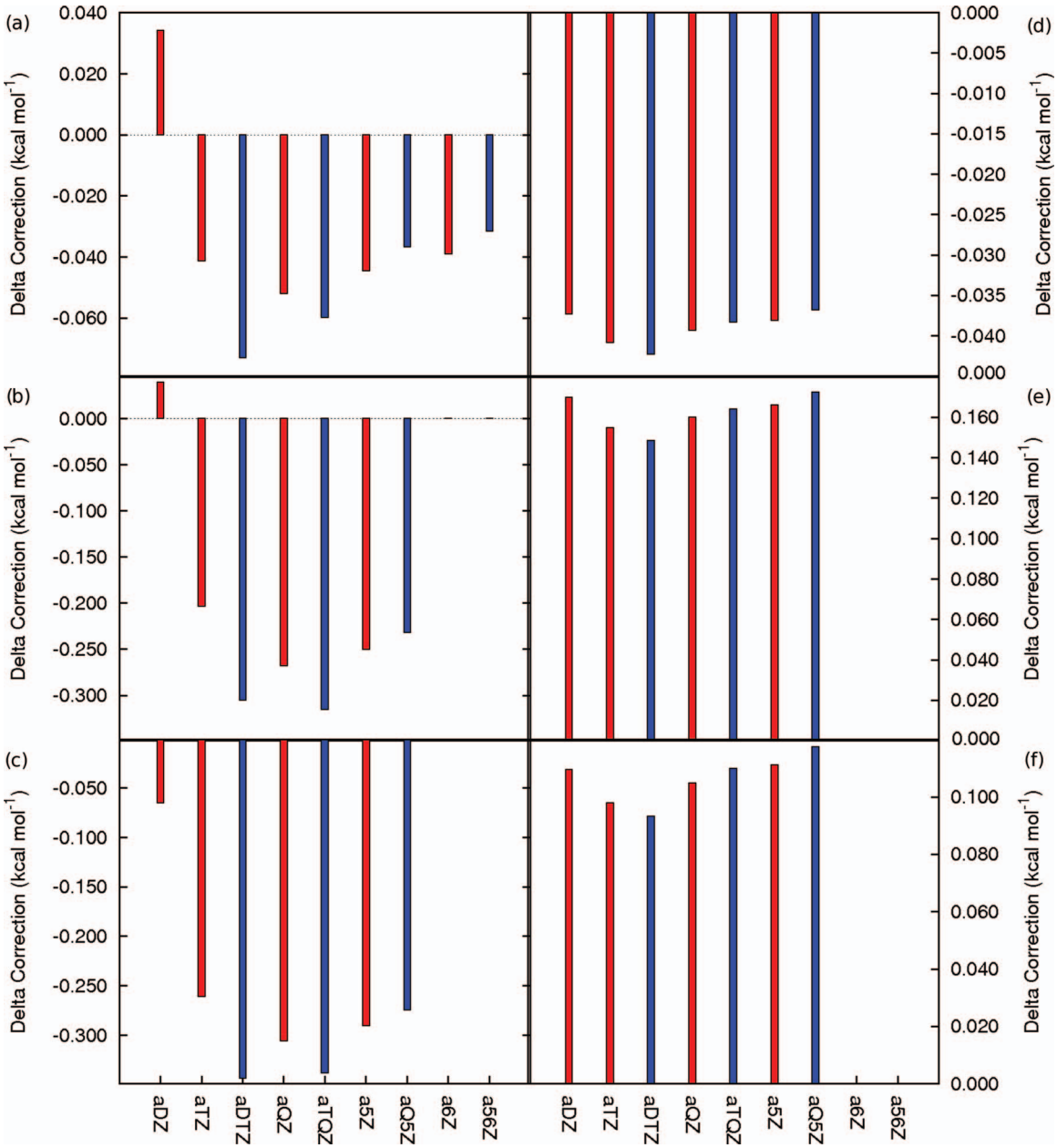


FIG. 1. Basis set convergence of $\delta_{MP2}^{CCSD(T)}$ for (a) water dimer, (b) formic acid dimer, (c) formamide dimer, (d) methane dimer, (e) ethene · ethine complex, and (f) ethene dimer. Hierarchical Dunning basis sets aDZ–a6Z (red bars) as well as their two-point Helgaker extrapolations (blue bars) are plotted, showing poor reliability of double-$\zeta$ results. All systems exhibit a "turnover" basis before which CBS estimates are unreliable and after which the term converges monotonically.

double-$\zeta$ basis sets (which work reasonably well for $\delta_{MP2}^{CCSD(T)}$ corrections in dispersion-dominated or mixed complexes) may perform better.

Methane dimer was investigated to see if similar qualitative basis set effects could be seen in dispersion bound complexes. Figure 1(d) illustrates the $\delta_{MP2}^{CCSD(T)}$ correction for the Dunning basis set series for methane dimer. Again, we see non-monotonically converging $\delta_{MP2}^{CCSD(T)}$ correction, but now the magnitude of the difference between CBS(aDZ,aTZ) and either aQZ or CBS(aQZ,a5Z) is 0.003 and 0.005 kcal mol$^{-1}$, respectively. While such errors are negligible, CBS(aDZ,aTZ) is still not recommended in place of single-basis $\delta_{MP2}^{CCSD(T)}$ because of the consistent over-correction. Ethene · ethine was investigated as a representative from the mixed category of the S22 test set. Figure 1(e) once again shows that the $\delta_{MP2}^{CCSD(T)}$ correction does not converge monotonically, instead featuring a turning point at aTZ. This leads CBS(aDZ,aTZ) extrapolation to overestimate the $\delta_{MP2}^{CCSD(T)}$ correction by 0.020 kcal mol$^{-1}$ compared to the CBS(aQZ,a5Z) value. Again, these errors are not very large, but neither is the magnitude of the $\delta_{MP2}^{CCSD(T)}$ correction. As the size of the $\delta_{MP2}^{CCSD(T)}$ correction increases for other van der Waals dimers, this error should increase as well, and using aTZ for $\delta_{MP2}^{CCSD(T)}$ instead of CBS(aDZ,aTZ) may offer a more noticeable improvement.

TABLE I. Benchmark interaction energies (kcal mol$^{-1}$) for the S22B database with references from which each component was taken.

| | Complex | Benchmark IE | Level of theory |
|---|---|---|---|
| 1 | HB ammonia dimer, $C_{2h}$ | − 3.133 | CCSD(T)/CBS(aQZ,a5Z)[d] |
| 2 | HB water dimer, $C_s$ | − 4.989 | CCSD(T)/CBS(a5Z,a6Z)[d] |
| 3 | HB formic acid dimer, $C_{2h}$ | − 18.753 | CCSD(T)/CBS(aQZ,a5Z)[d] |
| 4 | HB formamide dimer, $C_{2h}$ | − 16.062 | CCSD(T)/CBS(aQZ,a5Z)[d] |
| 5 | HB hydrogen-bonded uracil dimer, $C_{2h}$ | − 20.641 | MP2/CBS(haQZ-ha5Z)[b] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |
| 6 | HB 2-pyridone · 2-aminopyridine, $C_1$ | − 16.934 | MP2/CBS(haQZ-ha5Z)[b] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |
| 7 | HB adenine · thymine WC, $C_1$ | − 16.660 | MP2/CBS(haQZ-ha5Z)[b] + $\delta_{MP2}^{CCSD(T)}$/aTZ[c] |
| 8 | DD methane dimer, $D_{3d}$ | − 0.527 | CCSD(T)/CBS(aQZ,a5Z)[d] |
| 9 | DD ethene dimer, $D_{2d}$ | − 1.472 | CCSD(T)/CBS(aQZ,a5Z)[d] |
| 10 | DD benzene · methane, $C_3$ | − 1.448 | MP2/CBS(aQZ-a5Z)[a] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |
| 11 | DD parallel displaced benzene dimer, $C_{2h}$ | − 2.654 | MP2/CBS(aQZ-a5Z)[a] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |
| 12 | DD pyrazine dimer, $C_s$ | − 4.255 | MP2/CBS(aQZ-a5Z)[a] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |
| 13 | MX stacked uracil dimer, $C_2$ | − 9.805 | MP2/CBS(haQZ-ha5Z)[b] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |
| 14 | DD stacked indole · benzene, $C_1$ | − 4.524 | MP2/CBS(aQZ-a5Z)[a] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |
| 15 | MX stacked adenine · thymine, $C_1$ | − 11.730 | MP2/CBS(haQZ-ha5Z)[b] + $\delta_{MP2}^{CCSD(T)}$/aTZ[c] |
| 16 | MX ethene · ethine, $C_{2v}$ | − 1.496 | CCSD(T)/CBS(aQZ,a5Z)[d] |
| 17 | MX benzene · water, $C_s$ | − 3.275 | MP2/CBS(aQZ-a5Z)[a] + $\delta_{MP2}^{CCSD(T)}$/aQZ[d] |
| 18 | MX benzene · ammonia, $C_s$ | − 2.312 | MP2/CBS(aQZ-a5Z)[a] + $\delta_{MP2}^{CCSD(T)}$/aQZ[d] |
| 19 | MX benzene · hydrogen cyanide, $C_s$ | − 4.541 | MP2/CBS(aQZ-a5Z)[a] + $\delta_{MP2}^{CCSD(T)}$/aQZ[d] |
| 20 | DD T-shaped benzene dimer, $C_{2v}$ | − 2.717 | MP2/CBS(aQZ-a5Z)[a] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |
| 21 | MX T-shaped indole · benzene, $C_1$ | − 5.627 | MP2/CBS(aTZ-aQZ)[a] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |
| 22 | MX phenol dimer, $C_1$ | − 7.097 | MP2/CBS(haQZ-ha5Z)[b] + $\delta_{MP2}^{CCSD(T)}$/aTZ+MB[a] |

[a]Podeszwa *et al*. (Ref. 38), MB indicates mid bond functions.
[b]Marchetti *et al*. (Ref. 55).
[c]Takatani *et al*. (Ref. 37).
[d]This work.

TABLE II. S22B benchmark interaction energies and differences (kcal mol$^{-1}$) with respect to Jurecka *et al*. (Ref. 11, original publication), Podeszwa *et al*. (Ref. 38, revision), and Takatani *et al*. (Ref. 37, S22A revision) literature values.

| Complex | Jurecka | Podeszwa | Takatani | S22B | \|Δ Jurecka \| | \|Δ Podeszwa \| | \|Δ Takatani \| |
|---|---|---|---|---|---|---|---|
| 1 | − 3.17 | − 3.145 | − 3.150 | − 3.133 | 0.037 | 0.012 | 0.037 |
| 2 | − 5.02 | − 5.004 | − 5.070 | − 4.989 | 0.031 | 0.015 | 0.031 |
| 3 | − 18.61 | − 18.751 | − 18.810 | − 18.753 | 0.143 | 0.002 | 0.047 |
| 4 | − 15.96 | − 16.063 | − 16.110 | − 16.062 | 0.101 | 0.002 | 0.059 |
| 5 | − 20.65 | − 20.643 | − 20.690 | − 20.641 | 0.009 | 0.002 | 0.049 |
| 6 | − 16.71 | − 16.938 | − 17.000 | − 16.934 | 0.224 | 0.004 | 0.066 |
| 7 | − 16.37 | − 16.555 | − 16.740 | − 16.660 | 0.290 | 0.105 | 0.080 |
| 8 | − 0.53 | − 0.530 | − 0.530 | − 0.527 | 0.003 | 0.003 | 0.003 |
| 9 | − 1.51 | − 1.483 | − 1.480 | − 1.472 | 0.039 | 0.012 | 0.029 |
| 10 | − 1.5 | − 1.448 | − 1.450 | − 1.448 | 0.052 | 0.000 | 0.002 |
| 11 | − 2.73 | − 2.654 | − 2.620 | − 2.654 | 0.076 | 0.000 | 0.034 |
| 12 | − 4.42 | − 4.255 | − 4.200 | − 4.255 | 0.165 | 0.000 | 0.055 |
| 13 | − 10.12 | − 9.783 | − 9.740 | − 9.805 | 0.315 | 0.022 | 0.065 |
| 14 | − 5.22 | − 4.524 | − 4.590 | − 4.524 | 0.696 | 0.000 | 0.066 |
| 15 | − 12.23 | − 11.857 | − 11.660 | − 11.730 | 0.500 | 0.127 | 0.070 |
| 16 | − 1.53 | − 1.503 | − 1.500 | − 1.496 | 0.034 | 0.007 | 0.014 |
| 17 | − 3.28 | − 3.280 | − 3.290 | − 3.275 | 0.005 | 0.005 | 0.015 |
| 18 | − 2.35 | − 2.320 | − 2.320 | − 2.312 | 0.038 | 0.008 | 0.008 |
| 19 | − 4.46 | − 4.540 | − 4.550 | − 4.541 | 0.081 | 0.001 | 0.009 |
| 20 | − 2.74 | − 2.717 | − 2.710 | − 2.717 | 0.023 | 0.000 | 0.007 |
| 21 | − 5.73 | − 5.627 | − 5.620 | − 5.627 | 0.103 | 0.000 | 0.007 |
| 22 | − 7.05 | − 7.097 | − 7.090 | − 7.097 | 0.047 | 0.000 | 0.007 |
| | | | | Average | 0.137 | 0.015 | 0.035 |
| | | | | Max | 0.696 | 0.127 | 0.080 |

### B. Impact on current benchmark sets for non-covalent interactions

From this basis set study on the $\delta_{MP2}^{CCSD(T)}$ correction, we reach two general conclusions: (1) Extrapolated corrections should be avoided unless one can ensure that the basis sets used are beyond the turning point (typically aQZ for hydrogen-bonded and aTZ for mixed and dispersion bound complexes). These are admittedly large basis sets for CCSD(T) computations, which argues against using extrapolation techniques for $\delta_{MP2}^{CCSD(T)}$ corrections in general. (2) aDZ often results in the wrong sign for $\delta_{MP2}^{CCSD(T)}$ corrections for small hydrogen-bonded complexes (S22-1 through S22-7), and larger basis sets should be used whenever possible. In light of these new findings, we have decided to revise some existing benchmark sets that used extrapolated $\delta_{MP2}^{CCSD(T)}$ values. The best estimates for interaction energies of the S22 complexes have been revised as described in Table I and are herein referred to as S22B. This new set removes any CBS extrapolated $\delta_{MP2}^{CCSD(T)}$ corrections and uses the largest underlying MP2/CBS as well as the largest $\delta_{MP2}^{CCSD(T)}$ correction available from the literature. We also provide larger basis set $\delta_{MP2}^{CCSD(T)}$ corrections or directly extrapolated CCSD(T)/CBS values (without using a $\delta_{MP2}^{CCSD(T)}$ term) for some of the smallest complexes. From Table II, we see noticeable deviations for the hydrogen-bonded complexes compared to the S22A benchmark values of Takatani *et al.*[37] that used CBS(aDZ,aTZ) $\delta_{MP2}^{CCSD(T)}$ corrections. We also note substantial differences for the adenine · thymine complexes compared to the work of Podeszwa *et al.*,[38] due to the latter's use of an aDZ $\delta_{MP2}^{CCSD(T)}$ correction. S22B benchmark values differ from those of Jurecka *et al.*,[11] Podeszwa *et al.*,[38] and Takatani *et al.*[37] by 0.137, 0.015, and 0.035 kcal mol$^{-1}$ on average, respectively. Maximum differences are 0.696, 0.127, and 0.080 kcal mol $^{-1}$, respectively.

The NBC10 (Refs. 3, 34, and 35) test set also generally utilized CBS(aDZ,aTZ) and CBS(haDZ,haTZ) $\delta_{MP2}^{CCSD(T)}$ corrections, so we updated these to use only the aTZ and haTZ $\delta_{MP2}^{CCSD(T)}$ corrections. The geometries were not changed. Supplementary Tables S2–S6 (see Ref. 51) reflect the new values and the change to the new $\delta_{MP2}^{CCSD(T)}$ corrections. This new benchmark will be referred to as NBC10A. We report a shift of 0.017 kcal mol$^{-1}$ on average across all complexes and a maximum difference of 0.060 kcal mol$^{-1}$ for parallel displaced (PD) benzene dimer at an intermolecular separation of 3.2 Å and a slip distance of 0.2 Å. We note the largest corrections are for stacked configurations on the repulsive wall.

Since the overestimation of $\delta_{MP2}^{CCSD(T)}$ corrections by CBS(aDZ,aTZ) extrapolation primarily affects hydrogen-bonded systems, we must also revise the HBC6 test set which consists of doubly hydrogen-bonded complexes. The only modification to this test set was the replacement of CBS(aDZ,aTZ) $\delta_{MP2}^{CCSD(T)}$ corrections with pure aTZ $\delta_{MP2}^{CCSD(T)}$ corrections for all complexes. Tables III–V reflect the new values and the magnitude of the change to the new $\delta_{MP2}^{CCSD(T)}$ corrections, along with an incrementation of the test set name to "HBC6A." We report a shift of 0.073 kcal mol$^{-1}$ on average across all complexes and a maximum difference of 0.257 kcal mol$^{-1}$ for formamidine dimer at an intermolecular separation of 3.4 Å. Similar to NBC10, we note the largest corrections are for geometries on the repulsive wall.

TABLE III. HBC6A benchmark interaction energies (using HF/aQZ, MP2/CBS(aTZ,aQZ) extrapolation, and $\delta_{MP2}^{CCSD(T)}$/aTZ components) and differences (kcal mol$^{-1}$) with respect to Thanthiriwatte *et al.* (Ref. 36, original publication using $\delta_{MP2}^{CCSD(T)}$/CBS(aDZ,aTZ) extrapolation); formic acid (FaOO) and formamide (FaON) dimers at various intermolecular distances (in Å).

| Complex | Original | HBC6A | $\|\Delta\|$ |
|---|---|---|---|
| FaOO–FaOO-3.4 | − 19.834 | − 19.627 | 0.207 |
| FaOO–FaOO-3.5 | − 20.027 | − 19.850 | 0.177 |
| FaOO–FaOO-3.6 | − 20.060 | − 19.910 | 0.150 |
| FaOO–FaOO-3.7 | − 19.776 | − 19.650 | 0.126 |
| FaOO–FaOO-3.8 | − 19.132 | − 19.027 | 0.105 |
| FaOO–FaOO-3.9 | − 18.161 | − 18.075 | 0.086 |
| FaOO–FaOO-4.0 | − 16.943 | − 16.873 | 0.070 |
| FaOO–FaOO-4.1 | − 15.574 | − 15.517 | 0.057 |
| FaOO–FaOO-4.2 | − 14.148 | − 14.100 | 0.048 |
| FaOO–FaOO-4.3 | − 12.736 | − 12.697 | 0.039 |
| FaOO–FaOO-4.4 | − 11.392 | − 11.360 | 0.032 |
| FaOO–FaOO-4.6 | − 9.014 | − 8.990 | 0.024 |
| FaOO–FaOO-4.8 | − 7.091 | − 7.074 | 0.017 |
| FaOO–FaOO-5.0 | − 5.590 | − 5.577 | 0.013 |
| FaOO–FaOO-5.4 | − 3.548 | − 3.539 | 0.009 |
| FaOO–FaOO-5.8 | − 2.325 | − 2.323 | 0.002 |
| FaOO–FaOO-6.4 | − 1.320 | − 1.320 | 0.000 |
| FaOO–FaOO-7.0 | − 0.801 | − 0.802 | 0.001 |
| FaOO–FaOO-8.0 | − 0.394 | − 0.397 | 0.003 |
| FaOO–FaOO-10.0 | − 0.132 | − 0.135 | 0.003 |
| FaON–FaON-3.4 | − 6.726 | − 6.556 | 0.170 |
| FaON–FaON-3.5 | − 10.191 | − 10.027 | 0.164 |
| FaON–FaON-3.6 | − 12.781 | − 12.628 | 0.153 |
| FaON–FaON-3.7 | − 14.667 | − 14.529 | 0.138 |
| FaON–FaON-3.8 | − 15.919 | − 15.796 | 0.123 |
| FaON–FaON-3.9 | − 16.582 | − 16.475 | 0.107 |
| FaON–FaON-4.0 | − 16.714 | − 16.622 | 0.092 |
| FaON–FaON-4.1 | − 16.391 | − 16.313 | 0.078 |
| FaON–FaON-4.2 | − 15.713 | − 15.647 | 0.066 |
| FaON–FaON-4.3 | − 14.790 | − 14.735 | 0.055 |
| FaON–FaON-4.4 | − 13.723 | − 13.678 | 0.045 |
| FaON–FaON-4.6 | − 11.480 | − 11.448 | 0.032 |
| FaON–FaON-4.8 | − 9.401 | − 9.379 | 0.022 |
| FaON–FaON-5.0 | − 7.642 | − 7.626 | 0.016 |
| FaON–FaON-5.4 | − 5.108 | − 5.097 | 0.011 |
| FaON–FaON-5.8 | − 3.537 | − 3.528 | 0.009 |
| FaON–FaON-6.4 | − 2.187 | − 2.181 | 0.006 |
| FaON–FaON-7.0 | − 1.448 | − 1.443 | 0.005 |
| FaON–FaON-8.0 | − 0.816 | − 0.813 | 0.003 |
| FaON–FaON-10.0 | − 0.340 | − 0.337 | 0.003 |

The last test set revised in this study is HSG,[12] which is a benchmark formed by dissecting the reaction site of a bound protein-drug complex (HIV-II protease-indinavir) into 21 pairs of chemical fragments that are not necessarily at equilibrium geometries. Again, here we only replace the CBS(haDZ,haTZ) $\delta_{MP2}^{CCSD(T)}$ correction with a haTZ $\delta_{MP2}^{CCSD(T)}$ correction. Table VI reflects the new values and corresponding shifts for HSG-A. We report an average change of 0.027 kcal mol$^{-1}$ and a maximum change of 0.102 kcal

TABLE IV. HBC6A benchmark interaction energies (using HF/aQZ, MP2/CBS(aTZ,aQZ) extrapolation, and $\delta_{MP2}^{CCSD(T)}$/aTZ components) and differences (kcal mol$^{-1}$) with respect to Thanthiriwatte *et al.* (Ref. 36, original publication using $\delta_{MP2}^{CCSD(T)}$/CBS(aDZ,aTZ) extrapolation); dimers of formic acid (FaOO), formamide (FaON), and formamidine (FaNN) at various intermolecular distances (in Å).

| Complex | Original | HBC6A | $\|\Delta\|$ |
|---|---|---|---|
| FaNN–FaNN-3.4 | − 8.987 | − 8.730 | 0.257 |
| FaNN–FaNN-3.5 | − 10.969 | − 10.725 | 0.244 |
| FaNN–FaNN-3.6 | − 12.693 | − 12.463 | 0.230 |
| FaNN–FaNN-3.7 | − 14.144 | − 13.932 | 0.212 |
| FaNN–FaNN-3.8 | − 15.287 | − 15.106 | 0.181 |
| FaNN–FaNN-3.9 | − 16.118 | − 15.950 | 0.168 |
| FaNN–FaNN-4.0 | − 16.587 | − 16.440 | 0.147 |
| FaNN–FaNN-4.1 | − 16.702 | − 16.575 | 0.127 |
| FaNN–FaNN-4.2 | − 16.452 | − 16.344 | 0.108 |
| FaNN–FaNN-4.3 | − 15.901 | − 15.811 | 0.090 |
| FaNN–FaNN-4.4 | − 15.102 | − 15.028 | 0.074 |
| FaNN–FaNN-4.6 | − 13.047 | − 12.999 | 0.048 |
| FaNN–FaNN-4.8 | − 10.810 | − 10.780 | 0.030 |
| FaNN–FaNN-5.0 | − 8.733 | − 8.715 | 0.018 |
| FaNN–FaNN-5.4 | − 5.539 | − 5.532 | 0.007 |
| FaNN–FaNN-5.8 | − 3.521 | − 3.517 | 0.004 |
| FaNN–FaNN-6.4 | − 1.861 | − 1.861 | 0.000 |
| FaNN–FaNN-7.0 | − 1.050 | − 1.051 | 0.001 |
| FaNN–FaNN-8.0 | − 0.463 | − 0.466 | 0.003 |
| FaNN–FaNN-10.0 | − 0.123 | − 0.127 | 0.004 |
| FaOO–FaON-3.4 | − 14.356 | − 14.164 | 0.192 |
| FaOO–FaON-3.5 | − 16.486 | − 16.312 | 0.174 |
| FaOO–FaON-3.6 | − 17.833 | − 17.679 | 0.154 |
| FaOO–FaON-3.7 | − 18.543 | − 18.409 | 0.134 |
| FaOO–FaON-3.8 | − 18.692 | − 18.578 | 0.114 |
| FaOO–FaON-3.9 | − 18.347 | − 18.250 | 0.097 |
| FaOO–FaON-4.0 | − 17.592 | − 17.512 | 0.080 |
| FaOO–FaON-4.1 | − 16.537 | − 16.471 | 0.066 |
| FaOO–FaON-4.2 | − 15.300 | − 15.245 | 0.055 |
| FaOO–FaON-4.3 | − 13.989 | − 13.944 | 0.045 |
| FaOO–FaON-4.4 | − 12.684 | − 12.647 | 0.037 |
| FaOO–FaON-4.6 | − 10.274 | − 10.248 | 0.026 |
| FaOO–FaON-4.8 | − 8.245 | − 8.227 | 0.018 |
| FaOO–FaON-5.0 | − 6.613 | − 6.597 | 0.016 |
| FaOO–FaON-5.4 | − 4.330 | − 4.321 | 0.009 |
| FaOO–FaON-5.8 | − 2.935 | − 2.931 | 0.004 |
| FaOO–FaON-6.4 | − 1.753 | − 1.751 | 0.002 |
| FaOO–FaON-7.0 | − 1.121 | − 1.119 | 0.002 |
| FaOO–FaON-8.0 | − 0.598 | − 0.597 | 0.001 |
| FaOO–FaON-10.0 | − 0.227 | − 0.228 | 0.001 |

TABLE V. HBC6A benchmark interaction energies (using HF/aQZ, MP2/CBS(aTZ,aQZ) extrapolation, and $\delta_{MP2}^{CCSD(T)}$/aTZ components) and differences (kcal mol$^{-1}$) with respect to Thanthiriwatte *et al.* (Ref. 36, original publication using $\delta_{MP2}^{CCSD(T)}$/CBS(aDZ,aTZ) extrapolation); dimers of formamidine (FaNN) with formamide (FaON) and formic acid (FaOO) at various intermolecular distances (in Å).

| Complex | Original | HBC6A | $\|\Delta\|$ |
|---|---|---|---|
| FaON–FaNN-3.4 | − 8.239 | − 8.021 | 0.218 |
| FaON–FaNN-3.5 | − 10.918 | − 10.711 | 0.207 |
| FaON–FaNN-3.6 | − 13.055 | − 12.862 | 0.193 |
| FaON–FaNN-3.7 | − 14.717 | − 14.539 | 0.178 |
| FaON–FaNN-3.8 | − 15.921 | − 15.763 | 0.158 |
| FaON–FaNN-3.9 | − 16.672 | − 16.532 | 0.140 |
| FaON–FaNN-4.0 | − 16.977 | − 16.856 | 0.121 |
| FaON–FaNN-4.1 | − 16.865 | − 16.760 | 0.105 |
| FaON–FaNN-4.2 | − 16.390 | − 16.301 | 0.089 |
| FaON–FaNN-4.3 | − 15.631 | − 15.557 | 0.074 |
| FaON–FaNN-4.4 | − 14.676 | − 14.614 | 0.062 |
| FaON–FaNN-4.6 | − 12.490 | − 12.448 | 0.042 |
| FaON–FaNN-4.8 | − 10.304 | − 10.277 | 0.027 |
| FaON–FaNN-5.0 | − 8.362 | − 8.341 | 0.021 |
| FaON–FaNN-5.4 | − 5.445 | − 5.434 | 0.011 |
| FaON–FaNN-5.8 | − 3.617 | − 3.609 | 0.008 |
| FaON–FaNN-6.4 | − 2.087 | − 2.082 | 0.005 |
| FaON–FaNN-7.0 | − 1.295 | − 1.292 | 0.003 |
| FaON–FaNN-8.0 | − 0.663 | − 0.661 | 0.002 |
| FaON–FaNN-10.0 | − 0.237 | − 0.237 | 0.000 |
| FaOO–FaNN-3.6 | − 26.289 | − 26.064 | 0.225 |
| FaOO–FaNN-3.7 | − 24.035 | − 23.841 | 0.194 |
| FaOO–FaNN-3.8 | − 23.017 | − 22.850 | 0.167 |
| FaOO–FaNN-3.9 | − 22.133 | − 21.990 | 0.143 |
| FaOO–FaNN-4.0 | − 21.122 | − 21.002 | 0.120 |
| FaOO–FaNN-4.1 | − 19.920 | − 19.819 | 0.101 |
| FaOO–FaNN-4.2 | − 18.544 | − 18.461 | 0.083 |
| FaOO–FaNN-4.3 | − 17.056 | − 16.988 | 0.068 |
| FaOO–FaNN-4.4 | − 15.526 | − 15.471 | 0.055 |
| FaOO–FaNN-4.6 | − 12.583 | − 12.546 | 0.037 |
| FaOO–FaNN-4.8 | − 10.031 | − 10.006 | 0.025 |
| FaOO–FaNN-5.0 | − 7.960 | − 7.942 | 0.018 |
| FaOO–FaNN-5.4 | − 5.069 | − 5.058 | 0.011 |
| FaOO–FaNN-5.8 | − 3.336 | − 3.328 | 0.008 |
| FaOO–FaNN-6.4 | − 1.906 | − 1.900 | 0.006 |
| FaOO–FaNN-7.0 | − 1.170 | − 1.166 | 0.004 |
| FaOO–FaNN-8.0 | − 0.587 | − 0.584 | 0.003 |
| FaOO–FaNN-10.0 | − 0.202 | − 0.200 | 0.002 |

mol$^{-1}$. The small average change is due to there being only a few hydrogen-bonded complexes in the HSG test set.

The remaining errors in these benchmark test sets are anticipated to be the following: basis set incompleteness error (BSIE) of the MP2/CBS, BSIE of the $\delta_{MP2}^{CCSD(T)}$ correction, core-valence correction, and higher-order excitation corrections. To examine these sources of error, we look at each for the S22 benchmark set. By comparing MP2/CBS(aTZ,aQZ) to MP2/CBS(aQZ,a5Z), we estimate the BSIE of the MP2/CBS to have an average percent error of 0.10% (maximum 0.22% for PD benzene dimer). We estimate the upper bound of the BSIE of the $\delta_{MP2}^{CCSD(T)}$ by comparing $\delta_{MP2}^{CCSD(T)}$/aDZ and $\delta_{MP2}^{CCSD(T)}$/aTZ. This approach reveals an average percent error of 0.60% (maximum of 3.01% for PD benzene dimer). To understand the remaining BSIE better, we looked at both the counterpoise corrected and non-counterpoise corrected $\delta_{MP2}^{CCSD(T)}$ corrections (see supplementary Figures S1 and S2 given in Ref. 51). These figures reveal that there is still significant BSSE with the aTZ basis and that an uncorrected $\delta_{MP2}^{CCSD(T)}$ should therefore be used with caution (all our proposed benchmarks use the counterpoise corrected $\delta_{MP2}^{CCSD(T)}$). The core-valence correction was estimated by Podeszwa *et al.*[38] to be on the order of 0.1% and no larger than 0.5% for all molecules in the S22 test set. Because of a lack of detailed studies of higher-order corrections (with an adequate basis set) in dispersion dominated complexes,

TABLE VI. HSG-A benchmark interaction energies (using HF/aQZ, MP2/CBS(aTZ,aQZ) extrapolation, and $\delta_{MP2}^{CCSD(T)}$/haTZ components) and differences (kcal mol$^{-1}$) with respect to Faver *et al.* (Ref. 12 values, using $\delta_{MP2}^{CCSD(T)}$/CBS(haDZ,haTZ) extrapolation).

| | Complex | Original | HSG-A | $\|\Delta\|$ |
|---|---|---|---|---|
| 1 | ala29-big | −0.519 | −0.518 | 0.001 |
| 2 | ala128-small | −2.181 | −2.283 | 0.102 |
| 3 | arg8 | −2.451 | −2.478 | 0.027 |
| 4 | ash26-asp125 | −16.445 | −16.526 | 0.081 |
| 5 | asp129-big | −18.984 | −19.076 | 0.092 |
| 6 | asp130 | −6.009 | −5.998 | 0.011 |
| 7 | gly28-big | −3.301 | −3.308 | 0.007 |
| 8 | gly50-ring-big | −0.554 | −0.581 | 0.027 |
| 9 | gly50-v1 | −5.038 | −5.066 | 0.028 |
| 10 | gly127 | −7.532 | −7.509 | 0.023 |
| 11 | gly148 | −6.279 | −6.274 | 0.005 |
| 12 | ile48-big | 0.305 | 0.302 | 0.003 |
| 13 | ile147 | −2.087 | −2.103 | 0.016 |
| 14 | ile150-big | −1.376 | −1.378 | 0.002 |
| 15 | ile184 | −0.853 | −0.856 | 0.003 |
| 16 | leu23-big | −1.097 | −1.100 | 0.003 |
| 17 | pro181 | −1.504 | −1.534 | 0.030 |
| 18 | val33-big | −0.473 | −0.472 | 0.001 |
| 19 | val83 | −1.569 | −1.598 | 0.029 |
| 20 | val132 | 0.391 | 0.378 | 0.013 |
| 21 | wat200 | −9.486 | −9.538 | 0.052 |
| | | | Average | 0.027 |
| | | | Max | 0.102 |

post-triples corrections are hard to quantify. Hopkins *et al.*[52] estimated $\delta_{CCSD(T)}^{CCSD(TQ)}$ corrections to be approximately a tenth the magnitude of $\delta_{MP2}^{CCSD(T)}$ (with the same sign), but these were for relatively small systems. Pitonak *et al.*[22] reported quadruple excitation corrections for benzene dimer to be 0.04 kcal mol$^{-1}$ (1.72%), but this study used a relatively small 6-31G*(0.25) basis set. A more detailed study of higher-order corrections with adequate basis sets is required before giving bounds on this error.

### C. Small basis set $\delta_{MP2}^{CCSD(T)}$ corrections

Having established that benchmark-quality $\delta_{MP2}^{CCSD(T)}$ terms generally require triple-$\zeta$ basis sets to be truly robust, we examined the performance of small double-$\zeta$ basis sets often present in the literature. Table S1 in the supplementary material[51] presents our best estimates of the $\delta_{MP2}^{CCSD(T)}$ correction for the S22 test set, along with estimates of this correction evaluated in various double-$\zeta$ basis sets. Mean absolute deviations are presented in Figure 2(a). It is clear that the original 6-31G* basis should not be used for these types of computations, as the optimized version 6-31G*(0.25) significantly outperforms it for the same computational cost. 6-31G*(0.25) performs relatively well for hydrogen-bonded complexes (MAD is 0.09 kcal mol$^{-1}$), but the error becomes somewhat larger than desirable for mixed complexes (0.15 kcal mol$^{-1}$) and dispersion-dominated complexes (MAD 0.20 kcal mol$^{-1}$ and MAPD 6.8%). Figure 2 also illustrates the importance of diffuse functions, as progressing

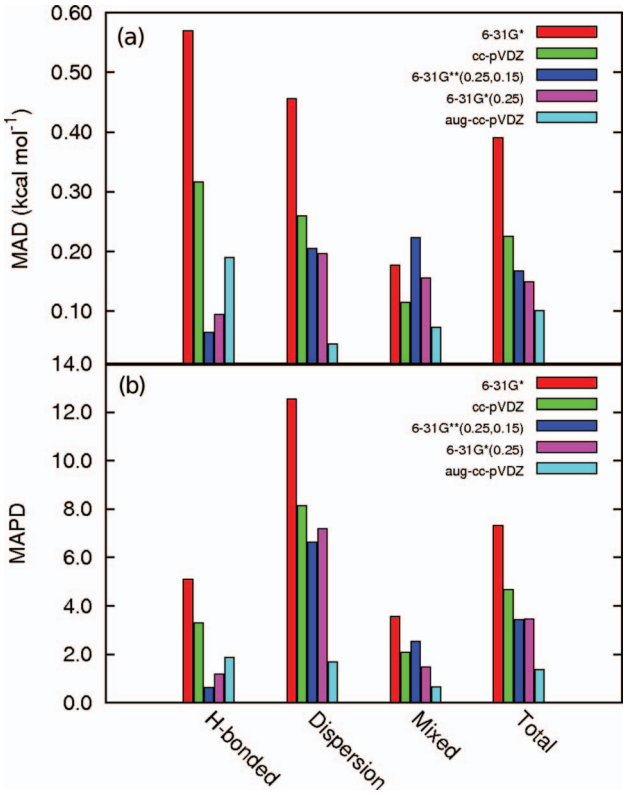


FIG. 2. Performance of double-$\zeta$ basis sets for the $\delta_{MP2}^{CCSD(T)}$ correction. For the S22 test set, (a) mean absolute deviations and (b) mean absolute percent deviations are assessed in relation to S22B benchmark CCSD(T)/CBS values. While the specially modified 6-31G*(0.25) and 6-31G**(0.25,0.15) basis sets yield low errors for hydrogen-bonding, they are significantly worse for dispersion-bound systems, and only aug-cc-pVDZ reaches an overall MAD for S22B of less than 0.1 kcal mol$^{-1}$.

from cc-pVDZ to aDZ reduces the MAD from 0.22 to 0.10 kcal mol$^{-1}$. Such augmented basis sets are especially important for dispersion bound complexes, reducing the MAD from 0.26 to 0.05 kcal mol$^{-1}$. Unfortunately, even aDZ is not an adequate basis set for high-quality $\delta_{MP2}^{CCSD(T)}$ corrections for all binding types; it produces a MAD of 0.18 kcal mol$^{-1}$ for hydrogen-bonded systems. Overall, Figure 2 confirms that none of the small double-$\zeta$ basis sets can produce an acceptable level of error for benchmark-quality interaction energies across all binding types, although with judicious choices, some may be sufficient for application studies or narrowly defined benchmarking tasks.

### D. Explicitly correlated $\delta_{MP2}^{CCSD(T)}$ corrections

We have shown that aDZ can lead to significant error in $\delta_{MP2}^{CCSD(T)}$ in some cases, so here we wanted to test how well this basis set could perform in an explicitly correlated framework. In Figure 3, the method labeled "MP2/CBS + δF12/aDZ" utilizes explicitly correlated $\delta_{MP2-F12}^{CCSD(T^*)-F12}$ corrections in an aDZ basis. MP2/CBS + δF12b/aDZ achieves 0.06 kcal mol$^{-1}$ MAD overall and 0.12, 0.02, and 0.04 for hydrogen bonding, dispersion, and mixed bonding, respectively. While this method performs very well, it does not significantly improve upon the CCSD(T**)-F12b/aDZ

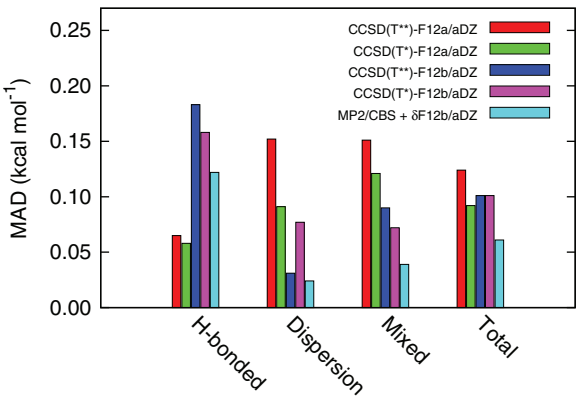


FIG. 3. Mean absolute deviation (MAD) for various explicitly correlated methods using the aug-cc-pVDZ basis versus S22B benchmark CCSD(T)/CBS values. The method labeled "MP2/CBS + $\delta$F12b/aDZ" is included to show the best performance of an estimated CCSD(T) approach that uses a MP2/CBS and a $\delta_{\mathrm{MP2-F12}}^{\mathrm{CCSD(T)-F12}}$ correction in an aug-cc-pVDZ basis.

approach itself. Apparently, the explicit correlation terms are so effective that CCSD(T**)-F12b/aDZ does not need to be mixed with MP2/CBS estimates to account for basis set effects. In particular, MP2/CBS + $\delta$F12/aDZ does not alleviate the maximum error incurred for hydrogen-bound complexes (0.23 kcal mol$^{-1}$ for formic acid dimer). Hence, focal-point schemes using CCSD(T)-F12 to evaluate $\delta_{\mathrm{MP2-F12}}^{\mathrm{CCSD(T^*)-F12}}$ corrections do not seem to offer a large advantage over the underlying CCSD(T)-F12 in a modest basis.

## IV. CONCLUSIONS

This work examines the error incurred by employing polarized double-$\zeta$ basis sets for the $\delta_{\mathrm{MP2}}^{\mathrm{CCSD(T)}}$ portion of a focal-point estimate of CCSD(T)/CBS interaction energies for non-covalent complexes. The error in the coupled-cluster correction for a given basis set varies according to the non-covalent bonding motif. Although polarized double-$\zeta$ basis sets generally yield adequate estimates of $\delta_{\mathrm{MP2}}^{\mathrm{CCSD(T)}}$, especially if diffuse functions are included (or if the $d$ exponent is made more diffuse), nevertheless the errors (MAD of 0.39, 0.15, 0.17, 0.23, and 0.10 for the 6-31G*, 6-31G*(0.25), 6-31G**(0.25,0.15), cc-pVDZ, and aDZ basis sets for the S22 molecules) are too large for the resulting CCSD(T)/CBS values to be of true benchmark quality, given that several approximate methods are now capable of reproducing benchmark interaction energies within a few tenths of one kcal mol$^{-1}$. Hence, the remaining errors in $\delta_{\mathrm{MP2}}^{\mathrm{CCSD(T)}}$ should be taken into consideration when comparing new methods against benchmark sets that utilize $\delta_{\mathrm{MP2}}^{\mathrm{CCSD(T)}}$ corrections with such basis sets. By examining the coupled-cluster correction in progressively higher basis sets, up to a6Z, a characteristic turning point was found, after which the quantity converges monotonically and before which CBS extrapolations are unreliable. Particularly, CBS extrapolated $\delta_{\mathrm{MP2}}^{\mathrm{CCSD(T)}}$ corrections should not be used for hydrogen-bonded complexes when employing aDZ and aTZ basis sets. We recommend simply using the largest single basis set affordable. We report revised benchmark values for the S22, NBC10, HBC6, and HSG test sets based on lessons learned in this work. Cartesian coordinates and revised interaction energies for these four test sets are available as supplementary material.[51] The recent revision[53] of the S66 test set[54] used haDZ,haTZ extrapolated $\delta_{\mathrm{MP2}}^{\mathrm{CCSD(T)}}$ corrections; for those systems, just as for these, we expect the extrapolation procedure to introduce small errors (on the order of a few hundredths of one kcal mol$^{-1}$, perhaps more for any systems with double hydrogen bonds) relative to the true $\delta_{\mathrm{MP2}}^{\mathrm{CCSD(T)}}$/CBS values. We observe remarkable performance by the explicitly correlated methods CCSD(T**)-F12a/b, even with a modest aDZ basis set, yielding an MAD of only 0.1 kcal mol$^{-1}$ over the S22B test set. Such small errors mean that we need benchmark CCSD(T)/CBS values that are at least this precisely known, highlighting the value of the revised benchmark energies proposed here.

## ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation (NSF) (Grant No. CHE-1011360). The computer resources of the Center for Computational Molecular Science and Technology are funded through a National Science Foundation (NSF) CRIF Award (CHE-0946869).

[1] A. L. L. East and W. D. Allen, J. Chem. Phys. **99**, 4638 (1993).
[2] A. G. Csaszar, W. D. Allen, and H. F. Schaefer, J. Chem. Phys. **108**, 9751 (1998).
[3] C. D. Sherrill, T. Takatani, and E. G. Hohenstein, J. Phys. Chem. A **113**, 10146 (2009).
[4] M. O. Sinnokrot, E. F. Valeev, and C. D. Sherrill, J. Am. Chem. Soc. **124**, 10887 (2002).
[5] P. Jurečka and P. Hobza, Chem. Phys. Lett. **365**, 89 (2002).
[6] P. Hobza and J. Šponer, J. Am. Chem. Soc. **124**, 11802 (2002).
[7] M. O. Sinnokrot and C. D. Sherrill, J. Phys. Chem. A **110**, 10656 (2006).
[8] M. Pitoňák, T. Janowski, P. Neogrády, P. Pulay, and P. Hobza, J. Chem. Theory Comput. **5**, 1761 (2009).
[9] P. Jurecka and P. Hobza, Chem. Phys. Lett. **365**, 89 (2002).
[10] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, and K. Tanabe, J. Am. Chem. Soc. **124**, 104 (2002).
[11] P. Jurečka, J. Šponer, J. Černý, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).
[12] J. C. Faver, M. L. Benson , X. A. He, B. P. Robers, B. Wang, M. S. Marshall, M. R. Kennedy, C. D. Sherrill, and K. M. Merz, J. Chem. Theory Comput. **7**, 790 (2011).
[13] L. M. J. Kroon-Batenburg and F. B. van Duijneveldt, J. Mol. Struct. **22**, 185 (1985).
[14] P. Hobza, A. Mehlhorn, P. Carsky, and R. Zahradnik, J. Mol. Struct. **31**, 387 (1986).
[15] P. Hobza and R. Zahradnik, Chem. Rev. **88**, 871 (1988).
[16] P. Hobza, J. Sponer, and M. Polasek, J. Am. Chem. Soc. **117**, 792 (1995).
[17] A. D. Boese, J. M. L. Martin, and W. Klopper, J. Phys. Chem. **111**, 11122 (2007).
[18] S. K. Min, E. C. Lee, H. M. Lee, D. Y. Kim, D. Kim, and K. S. Kim, J. Comput. Chem. **29**, 1208 (2007).
[19] M. O. Sinnokrot and C. D. Sherrill, J. Phys. Chem. A **108**, 10200 (2004).
[20] R. Podeszwa, R. Bukowski, and K. Szalewicz, J. Phys. Chem. A **110**, 10345 (2006).
[21] T. Janowski and P. Pulay, Chem. Phys. Lett. **447**, 27 (2007).
[22] M. Pitoňák, P. Neogrády, J. Řezáč, P. Jurečka, M. Urban, and P. Hobza, J. Chem. Theory Comput. **4**, 1829 (2008).
[23] J. Gräfenstein and D. Cremer, J. Chem. Phys. **130**, 124105 (2009).
[24] L. Gráfová, M. Pitoňák, J. Řezáč, and P. Hobza, J. Chem. Theory Comput. **6**, 2365 (2010).
[25] L. A. Burns, Á. Vázquez-Mayagoitia, B. G. Sumpter, and C. D. Sherrill, J. Chem. Phys. **134**, 084107 (2011).
[26] H.-J. Werner and F. R. Manby, J. Chem. Phys. **124**, 054114 (2006).

[27]F. R. Manby, H.-J. Werner, T. B. Adler, and A. J. May, J. Chem. Phys. **124**, 094103 (2006).
[28]H.-J. Werner, J. Chem. Phys. **124**, 101103 (2008).
[29]H.-J. Werner, T. B. Adler, and F. R. Manby, J. Chem. Phys. **126**, 164102 (2007).
[30]T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[31]G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130**, 054104 (2009).
[32]S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).
[33]A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).
[34]T. Takatani and C. D. Sherrill, Phys. Chem. Chem. Phys. **9**, 6106 (2007).
[35]E. G. Hohenstein and C. D. Sherrill, J. Phys. Chem. A **113**, 878 (2009).
[36]K. S. Thanthiriwatte, E. G. Hohenstein, L. A. Burns, and C. D. Sherrill, J. Chem. Theory Comput. **7**, 88 (2011).
[37]T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, and C. D. Sherrill, J. Chem. Phys. **132**, 144104 (2010).
[38]R. Podeszwa, K. Patkowski, and K. Szalewicz, Phys. Chem. Chem. Phys. **12**, 5974 (2010).
[39]E. G. Hohenstein and C. D. Sherrill, "Wavefunction methods for noncovalent interactions," WIREs Comput. Mol. Sci. (in press).
[40]D. D. Nelson, G. T. Fraser, and W. Klemperer, Science **238**, 1670 (1987).
[41]E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci, and D. J. Nesbitt, Pure Appl. Chem. **83**, 1637 (2011).
[42]M. Pitonak, P. Neogrady, and P. Hobza, Phys. Chem. Chem. Phys. **12**, 1369 (2010).
[43]P. Jurečka, J. Šponer, and P. Hobza, J. Phys. Chem. B **108**, 5466 (2004).
[44]J. Pittner and P. Hobza, Chem. Phys. Lett. **390**, 496 (2004).
[45]I. Dabkowska, H. V. Gonzalez, P. Jurecka, and P. Hobza, J. Phys. Chem. A **109**, 1131 (2005).
[46]Y. Zhao and D. G. Truhlar, J. Phys. Chem. A **109**, 6624 (2005).
[47]J. Rejnek and P. Hobza, J. Phys. Chem. B **111**, 641 (2007).
[48]L. R. Rutledge and S. D. Westmore, J. Chem. Theory Comput. **4**, 1768 (2008).
[49]L. R. Rutledge, H. F. Durst, and S. D. Westmore, J. Chem. Theory Comput. **5**, 1400 (2009).
[50]O. Marchetti and H.-J. Werner, J. Phys. Chem. A **113**, 11580 (2009).
[51]See supplementary material at http://dx.doi.org/10.1063/1.3659142 for Figures S1–S2 showing the basis set convergence of counterpoise-corrected and uncorrected values of $\delta_{MP2}^{CCSD(T)}$ for the ethene dimer and water dimer, respectively; for Table S1 containing a comparison of $\delta_{MP2}^{CCSD(T)}$ corrections using various polarized double-$\zeta$ basis sets versus benchmark values; and for Tables S2–S9 detailing the reference energies of the NBC10A database. Text files with revised benchmark interaction energies for the S22B, NBC10A, HBC6A, and HSG-A databases are also available along with Cartesian coordinates for all test set members.
[52]B. W. Hopkins and G. S. Tschumper, J. Phys. Chem. A **108**, 2941 (2004).
[53]J. Řezác, K. E. Riley, and P. Hobza, "Extensions of the S66 data set: More accurate interaction energies and angular-displaced nonequilibrium geometries," J. Chem. Theory Comput. (in press).
[54]J. Řezác, K. E. Riley, and P. Hobza, J. Chem. Theory Comput. **7**, 2427 (2011).
[55]O. Marchetti and H.-J. Werner, Phys. Chem. Chem. Phys. **10**, 3400 (2008).